WPS4058

Learning through Monitoring:

Lessons from a Large Scale Nutrition Program in Madagascar [†]

Emanuela Galasso and Jeffrey Yau

Development Research Group
World Bank

Bureau of Economics
*Federal Trade Commission**

Abstract:

Monitoring data is generally collected as a by-product of the process of monitoring program implementation. Yet this rich source of data has not been exploited to assess the effectiveness of the program. In this paper we use detailed administered data from a large-scale, community-based nutrition program in Madagascar, to argue that this data can be used to estimate the *differential* effect of increased exposure to the program and study how these returns to exposure evolve over time. We find that the returns to exposure are positive: communities exposed for additional one (or two) years display on average lower malnutrition rates of around 7-9 percentage points. Moreover, we find that the returns are decreasing as time and duration increase, though they do not dissipate to zero. These results are consistent with the hypothesis that the returns to the program reflect learning effects from the intervention. Finally, the results show higher differential returns to the program in poorer areas and areas more vulnerable to diseases. These findings have important implications for how such programs should be scaled-up within a country.

Keywords: impact evaluation, duration, nutrition intervention, community-based program, large-scale programs, propensity-score, matching
JEL classification: I12, C14, C31

World Bank Policy Research Working Paper 4058, November 2006

[†] This research has been sponsored by the World Bank Research Grant "Community Nutrition: Evaluation of Impacts in Africa", a comparative meta-analysis on the effectiveness of community-based nutrition programs in Madagascar and Senegal. We thank Claudia Rokx, the project task manager for her enthusiastic support to the evaluation effort, Harold Alderman for useful discussions throughout the project, Rinku Murgai, Martin Ravallion and seminar participants at the NEUDC at Brown University, the World Bank and Virginia Tech. We are extremely grateful to the Seecaline project management unit, to Malalanirina Ratomaharo in particular, and to all the provincial directors for providing the monitoring data and sharing their vast knowledge of the program. We thank Bart Minten for sharing the data on the Commune Census. We wish to thank Quinghua Zhao for providing technical help in the data entry. The findings, interpretations and conclusions of this paper are those of the authors and should not be attributed to the World Bank, its Executive Directors, or countries they represent, the Federal Trade Commission or any individual commissioner.
Correspondence by emails to egalasso@worldbank.org or yyau@ftc.gov

Introduction

Monitoring and supervision are intrinsic aspects of the management of any social programs. Monitoring over time how the program performs in the field allows administrators to understand whether the project is implemented according to its design. An essential data tool underlying the monitoring activities is usually referred to as 'monitoring data', or MIS (monitoring and information system), which is a subset of the administrative data that is specifically linked to activities of a specific program. The data contained in the monitoring system generally produces simple descriptive statistics that, at different level of aggregation, provide feedback to the implementing agents to address eventual problems. Administrative data have been increasingly used as a tool to evaluate program performance of public assistance programs in developed countries, but there is still very limited evidence to date that uses administrative data to evaluate social programs in developing countries and hardly any that uses monitoring data for this purpose. [1]

Why is it the case? The main reason constraining the use monitoring data for evaluation purposes is that, by design, it only collects information on participants, therefore making it impossible to estimate what would have happened in the absence of the program. Evaluating a public program requires data from which to construct participant groups and comparisons groups. Simple *reflexive* comparisons (before-after) using only data on participants would generally impose a very strong identification assumption (Heckman, Lalonde, Smith 1999): they implicitly assign all the observed change in the outcomes of interest to the program, and assume away the need to construct a counterfactual change in the absence of the program[2]. As such, evaluations using monitoring data would not allow quantifying the parameters of interest generally used in the evaluation literature, such as the average treatment effect or the average effect of treatment on the treated, both of which require information on non-participants.

Yet, monitoring data have inherent (but un-explored) features that make them an attractive tool for analytical and evaluation work. First, it covers the entire universe of the participating population, allowing the results based

[1] A notable exception is Ravallion (2000), who uses administrative data on the geographical allocation of expenditures of a workfare program in Argentina matched with a poverty map to implement a methodology to measure the latent differences in the average allocation of the program between poor and non-poor.
[2] An interesting variation to using reflexive comparisons is proposed by Piehl et al. (2003), who show that sufficiently long time-series data on participants can be used to test for structural breaks and identifying impacts.

on this data to generalize to the entire target population and to study how the same program might have different outcomes in different geographic locations depending on their initial conditions. Second, it is created as a natural by-product of program implementation, making it a readily-available, low-cost source of data. Finally, it is collected since program inception and therefore has a time dimension that would be worth exploring when assessing the effectiveness of a program over time.

In this paper we argue that monitoring data can be fruitfully analyzed in the context of assessing program impact to address a different set of questions, questions that are highly relevant for programs where the treatment impact crucially depends on duration of exposure to the program. Specifically, we examine the relationship between the duration of program exposure and nutrition outcomes. The returns to exposure is important from a policy perspective in programs where gains are cumulative over time and where (as in most cases) the program impacts are expected to be heterogeneous across socio-economic groups or areas in across the country. If the returns to exposure are positive and the program is gradually phased in within a country, it might pay to target sequentially groups/areas with the highest marginal return to exposure first, in order to maximize the total gains from the investment in the program.

When estimating the differential returns to program duration on child nutritional outcomes, we use a community-based nutrition intervention in Madagascar. The program (SEECALINE)[3] was initiated under World Bank support in 1999 and has been gradually scaled-up to cover about one-third of the total targeted population. Its gradual expansion provides substantial variation in program exposure. Its large scale feature allows us to exploit the richness of the monitoring data that was collected over time during its implementation.

We quantify the marginal effect of increasing duration to the program by comparing at any given point in time the nutritional outcomes of the "comparable" communities that experience different durations of program exposure. We call this parameter of interest *differential treatment effect (DTE).* Figure 1, which plots the outcome trends for sites joining the program at different points in time, illustrates this scenario. The vertical distance between these curves is the observed difference in outcomes between the longer and the shorter exposure sites. These vertical

[3] Seecaline stands for "*Surveillance et Education des Ecoles et des Communautés en matière d'ALImentation et de Nutrition Elargie*" (Surveillance and Education of Schools and Communities on Food and General Nutrition) and is among the few examples of African countries to embark in a nationwide, large-scale community-based nutrition program.

differences represent the parameters to be estimated in our empirical work at any given point in time. Indeed, there are very few papers that take into account the *duration in program implementation* as an important parameter of interest, with the exception being Behrman, Cheng and Todd (2004), Armecin et al (2005), and Gertler (2004). Based on household survey data, these papers suggest that the marginal impact of increased program exposure is positive *at a given point in time*. Obviously, a simple comparison among communities with different lengths of exposure is likely to produce biased estimates to the extent that the phase-in is not made at random. In order to account for this potential selection into different durations, we adopt an empirical strategy that was first proposed by Behrman, Cheng and Todd (2004) that only require the use of program participants. The attractiveness of applying these methods to the monitoring data is that the differential effect of exposure can be estimated *at different points in time*. As time progresses, the pattern of the differential effect can provide suggestive evidence on how the returns evolve over time.

Why should duration matter in the context of this type of program, or more generally in the context of social programs? Duration might simply reflect *composition effects*: as in any other program that is gradually phased-in, it takes time for the program to fully cover the target population. At the program inception in the community only a subset of the target population will have benefited from it. Over time, a larger fraction of life for these children is naturally going to be exposed to the program. Assuming a positive impact of the program, communities that are exposed for longer duration should therefore have on average better outcomes relative to areas where children have exposed for less. Under this assumption, the returns to exposure should be positive and decreasing over time. At the limit, when both sets of communities have their children fully exposed to the program, the differential effect should be converging to zero.

A complementary, though not exclusive, explanation depends on the nature of the intervention. Community-based nutrition programs target at children early in life (generally from zero to three years of age), and promote behavioral change in the context of existing resources, acting through improvement in the knowledge of child care and feeding practices[4]. In this context, duration might matter per se because of *learning effects*. These effects may

[4] Mothers are taught the importance of exclusive breastfeeding (at least for infants), the timing and composition of complementary feeding, and hygiene and nutrition practices, especially when the child is ill. The basic package of growth monitoring and promotion is usually accompanied by nutrition IEC and nutrition education, micro-nutrient supplementation to

arise from improved mothers' knowledge on feeding and general nutritional practices that results from their participation to the program (direct effect). They may also work as an indirect effect of increased overall learning by all mothers absorbing nutrition education messages as well as by community workers who become more efficient in their program implementation. In this setting, the differential effect of exposure should be positive, but not necessarily be eliminated over time. Whether these returns are increasing/constant/or decreasing over time will depend on the specific technological features of these learning effects in the children's health production.

Our first set of results indicate that communities with longer program exposure show significantly better nutritional outcomes, measured in terms of lower malnutrition rates. These results echo the findings from the existing literature that returns to program exposure are positive at a given point in time. For one extra year of exposure, the differential effect is declining with age, and it is statistically significant only for the 0-6 months old children but not for older children (7-12 and 13-36 months old). For two extra year of exposure, this "age-ranking" phenomenon disappears, and the returns to exposure are significant for all age groups. More specifically, communities exposed to the program for two extra years, where most of the target population has been covered by the program, display on average lower malnutrition rates by 7-9 percentage points for all age groups. The results on the age-ranking suggest that the program impact takes time to materialize, and the younger children stand to gain more because a larger fraction of their lives has been affected by the program.

Second, the differential effects of exposure are decreasing over time for all age groups, even accounting for seasonal effects. It is important to highlight the fact that the differential effects of exposure, however, do not disappear as duration increases. We can ignore the composition effect when focusing on the cohort of 0-6 months old, who are rapidly fully exposed in both sets of communities (short and long) after six months of observation. Even when fully exposed to the program, we find that the returns to exposure for this age group decrease over time but do not converge to zero. We interpret this finding as supportive evidence towards learning effects of the program.

After examining the overall effect of differential length of exposure, we investigate the extent to which these effects vary along several key environmental factors that affect children's health in the intervention areas. These

children and pregnant women. One important variation across countries is whether or not supplementary feeding of malnourished children is included.

factors are (i) whether or not an intervention area has been hit by cyclone; (ii) poverty rate; (iii) the length of the lean season; and (iv) accessibility to safe water. We exploit the large-scale aspect of the program and the large sample size of our monitoring data (representing the universe of participants' communities) to re-estimate the returns to exposure for different subgroups of the targeted population and analyze whether the return differs across the subgroups.

Our results show that the returns are higher in poorer areas, in areas that have been exposed to cyclones, and in areas that have a longer lean season. The fact that the results vary significantly according to these key characteristics of the program sites has an important implication for the potential expansion of the program within the current districts of intervention as well as to the new areas in the rest of the country.

The rest of the paper is structured as follow: next section reviews the relevant literature, and section 3 describes the program and the monitoring data in details, highlighting the phase-in aspect observed in the data and the large variation in both initial nutritional outcome and economic characteristics. After fully describing the program features and the data structure, we will discuss the empirical method that we employ to estimate the return to the program duration, and our results are presented in section 6. Section 7 concludes the paper.

2. Literature Review

There exist only a few papers that focus on the *duration in program implementation* as the primary parameter of interest. Behrman, Cheng and Todd (2004) studies the impact of a pre-school program in Bolivia and finds that the impact on early childhood development outcomes is observed among children who have been exposed to the program for more than one year, compares to those exposed for shorter durations. They develop empirical methods to estimate the differential effect of duration both using a survey sample of participants and non-participants as well as using a sample of participating children only. An analogous approach is followed by Armecin et al (2005), who find outcomes for young children that are increasing with duration of exposure using household data from an early childhood development program in the Philippines. Finally, Gertler (2004) shows that the improvements in terms of the incidence of illness of children in the Mexican anti-poverty program *PROGRESA* is significant only for those villages with two years of program exposure, relatively to those with less than one year.

These papers consistently suggest that the marginal impact of increased program exposure is positive *at a given point in time*. Our empirical analysis builds on this existing evidence by providing estimates on how the returns of exposure evolve as time and duration into the program increase.

In addition to the small number of papers that study the impact of program duration, there also exists very little empirical evidence on how effective in terms of reducing malnutrition these types of interventions are for *large-scale* programs (Gillespie, Mason, Martorell 1996, Heaver 2002a).[5] The quantitative evidences available on the impact of community-based, nutrition intervention and other direct nutritional programs mainly come from either case studies or small-scale programs.[6] The only large-scale program that includes a nutrition component and that has been rigorously evaluated is *PROGRESA* (now *Oportunidades*) implemented in Mexico.[7] Even in the context of a randomized design, Attanasio et al (2003) have documented sizable differences in both the distribution of background characteristics as well as the returns on those characteristics across poorest and non-poorest states within the same country. They show that household observable characteristics do not provide sufficient information to extrapolate the effects of the intervention from one set of regions to another. Therefore, results from case studies

---

5 Many evaluation exercises of large-scale nutrition programs have generally not been planned ex-ante in the sense of placing particular attention to carefully construct a set of potential comparable, control communities or populations. This is the case for the evaluation of the Tamil Nadu Integrated Nutrition project (TINP), which shows a decrease in underweight prevalence of 1.5 per cent per year in participating districts (Heaver 2002b). Comparing differential trends over time in intervention areas and non-intervention areas may over/under-estimate the true changes attributed to the intervention if the selection bias due to the purposive targeting is not adequately addressed (Pitt, Rosenzweig, Gibbons 1993). Another notable example is the famous growth monitoring and promotion program in Honduras (Atencion Integral a la Ninez, AIN, Plowman et al 2004): AIN communities were chosen to be relatively more disadvantaged. The evaluation approach employed was such that the matched comparison communities ended up being exposed by the program, so that the evaluation focused on *reflexive* comparisons (before-after).

6 The first large-scale, community-based nutrition programs were implemented in South Asia (Tamil Nadu Integrated Nutrition Project (TINP) and the Bangladesh Integrated Nutrition Project (BINP)), but they suffered from weak evaluation design that failed to disentangle the effect of the program itself from the effects of other programs and of the general welfare improvements arising from economic growth. Notable examples include the Iringa project in Tanzania (Gillespie et al 2003), the Indonesian Nutrition Development Program and the Weaning Program in Indonesia (Favin Griffiths 1999), the ANEP program in the Dominican Republic (USAID 1988), the Nutrition Communication Project in Mali (Ross 1997)

7 The program, however, does not focus exclusively on nutrition. It offers a conditional cash transfer, conditional the households being able to fulfill some behavioral requirements which are designed to improve, among other things, prenatal care, immunization, nutrition monitoring, and supplementation and participation in health and nutrition education. With a randomized design during the phase-in, the program has shown significant improvements (among other outcomes) in children exposure to illnesses, anemia and height (Gertler 2004). The extrapolation of the results to other settings, however, is difficult for various reasons. On one hand, its design of the program makes it hard to disentangle the effects that are attributed to the cash component and that are attributed to the behavioral change induced by the cash transfer. Community based programs do not have a cash-transfer component and use voluntary participation to change behavior. On the other hand, the estimated impacts may not generalize to poorer contexts outside Latin America, especially to South Asia and Africa, where the bulk of the number of malnourished children and the malnutrition rates are higher, and where the problems of food security are much more severe.

or small-scale programs, even if they are well-designed randomized field trials, can hardly be generalized to a much larger population. Our paper will exploit the large scale dimension of the monitoring data to document the extent to which the specific parameter of program impact (in our instance, returns to exposure) varies along some key socio-economic characteristics in the context of a large scale, nutrition intervention.

3. The Program Seecaline: Program Design and Phase-in

The program SEECALINE is a community-based nutrition program that started in 1999 and is gradually scaled up to cover more than half of the country's districts. The objective of the program is to improve the nutritional status of children under the age of three and of pregnant and lactating mothers in the targeted project areas. To maximize geographical coverage as well as to provide quality services on a large-scale, the program is contracted out to local NGOs for implementation (management, delivery, operations research and supervision) at the local level[8] where the services are delivered locally by a community nutrition worker (ACN), who is usually an elected woman from the targeted community. The ACN's activities are in turn supervised by the NGO[9].

The preventive approach: Although conventional wisdom suggests that the *quantity of food* is the key to prevent growth faltering and retardation, recent studies have argued that nutrition practices, feeding methods, and the protection from diseases are equally important (Allen, Gillespie 2001, Cebu Study Team 1992, WHO 2003). Seecaline adopts a preventative approach to prevent growth from faltering before it becomes too severe. As a community-based nutrition program, Seecaline revolves around a monthly growth monitoring activity as a focal point. The participating communities are mobilized towards raising the awareness of the problem of malnutrition and are taught and encouraged to improve hygiene, child care and nutrition practices. All children under the age of three and the pregnant and lactating women in a community are eligible to participate in all of the program activities. On a monthly basis, the ACN weighs all of the children under the age of three and provides counseling to the mothers regarding the nutritional status of their children indicated by the growth chart. The ACN may also visit a participating family if its child's growth chart shows no progress or if he/she misses a weighing sessions.

[8] Specifically, by local level, we mean "sites," which are defined to be geographically bound by a five kilometer radius. This setup is designed to minimize the (direct and indirect) transport cost incurred by the mothers to participate in the site activities as well as to minimize the ACN''s workload. In this paper, we will use "sites" and "communities" interchangeably.
[9] In the monthly situation meeting, the ACN and the NGO representatives discuss the problems of the participating community.

Besides growth monitoring, the program provides other activities. Eligible children are given micronutrient supplementation and de-worming. The promotion of behavioral change includes direct counseling to the lactating and pregnant women, nutritional education sessions and cooking demonstrations by the community nutrition worker who heavily emphasizes the proper weaning practices and prepares recipes that make use of available the food sources that are available in the local community. The community nutrition worker also carries out an annual census of all of children under the age of three and mobilizes the mothers to participate in the weighing and education sessions.

Seecaline is a large-scale, community-based nutrition program that was gradually phased-in over time. The rest of this section will describe in detail the phase-in process. This description provides a crucial guidance regarding the variables used to model whether a community belongs to a longer vs. shorter duration.

Selection and timing of participation of districts of intervention: Madagascar is administratively organized in six provinces, 111 districts, around 1,400 communes and 16,000 villages. Seecaline's activities started in 1999 in four provinces and expanded to all six provinces in 2000.[10] The program targets the most malnourished districts: based on a nationally representative anthropometric survey collected in 1997/98, only the districts (46 out of 111) that had an average malnutrition rates above the national average were selected for the intervention. The total number of sites to be opened in each district was chosen to reach a coverage rate of 50% of all children below three years of age. The 50% target was to be achieved gradually over time, with 10% coverage per year in the first two years of operation and 15% coverage per year in the third and forth years. In addition, ten rural districts affected by droughts and cyclones in the year 2000 were selected to participate.[11] The program expansion stopped at the end of 2001/beginning of 2002. The intervention currently includes about 3,600 project sites.[12]

Selection and timing of participation of communities within the intervention districts: The program expansion followed a sequential contractual engagement of NGOs across districts. NGO were selected based on a bidding

---

10 CNII is an extension of an earlier Bank pilot project originally targeted to two provinces (Antananarivo and Toliary).

11 Finally six additional urban districts that cover the province capitals were added in 2002, in the aftermath of a political crisis. Emergency sites were opened to alleviate the cost of the crisis on urban populations, and were subsequently made permanent. Our data is not available for such urban areas.

12 Note that we are using communities and program sites interchangeably. A site is identified by its geographical delimitation (of a radius of five kilometers, within the commune boundaries). A site generally comprises 1 to 3 villages.

workshop organized every year around August/September[13]. The selected NGOs were then assigned a given number of sites to be open in any given district, based on an estimate of the target population in the districts.[14]

The program in each province was initially advertised to all majors of the communes included within the district. Mayors would then organize a meeting with all the communities to make them aware of the program and encourage them to participate. A community, to be eligible to open a project site, had to accessible for most part of the year (by auto/motorbike, chariot or pirogue). Accessibility in Madagascar is a major constraint for service delivery, and the phasing in of the program reflected this constraint. The NGO, and more specifically the animator who is in charge of supervising and providing support to the community nutrition worker, need to able to reach the sites regularly by motorbike. Given the logistical constraints, the NGOs, in partnership with the communes, started on average with the most accessible communities first, and then expanded to the least accessible sites over time.

4. The Data

The monitoring data of the program consists of monthly detailed information at the site level. This information is originated from the registry and is kept by the community nutrition worker. Each registry contains information on the program activities, health outcomes, and the time when the site joins the program.

The monitoring data covers about 3,600 sites in the whole country (as of January 2003). The information spans from 1999 to the end of 2002, with information varying depending on the date of entry to the program. We will confine the analysis to the biggest four main provinces (Antananarivo, Fianarantsoa, Tulear and Tamatave), which represent 85% of the total target population, and for which there is substantial variation in the phase-in of the program. We will make use of the quarterly data in order to smooth out noises in the outcome indicators.

The registry for each site contains information about the monthly participation of registered mothers, the total number of children weighed in each month, and the number per age group who are classified as severely or moderately malnourished (as indicated by their position in the growth chart).[15] Moreover, for each community we

[13] Eligible NGOs had to have had at least one year of experience in the district of intervention.
[14] Each NGO could be assigned a maximum of 20 sites and had to hire a supervisor (*animateur*) to be assigned to each site.
[15] The registry contains also information on food supplementation (total quantity received, and the share of children receiving any food supplementation by age group), vitamin A supplementation (by age group) and on participation rates of mothers to nutritional education sessions and cooking demonstrations.

have information of the distance of the site to the center of the commune, as well as basic information on each partner NGO that is in charge of the site (such as size/personnel, whether it is local or national, their experience and sector of activity before joining Seecaline).

We complement the monitoring data with two additional sources of data. The first one is the Commune Census data, conducted in 2001 under a joint collaboration between Cornell University, the National Statistical Institute (INSTAT) and the agricultural research institute within the Ministry of Scientific Research (FOFIFA). The census contains detailed information on demographic and socio-economic characteristics of all communes in the country[16]. The most important aspect of this data is that it includes some key variables that capture the site selection process such as remoteness, main economic activities, and access to infrastructure. Second, we use commune level estimates[17] of poverty at the commune from the poverty map developed by Mistiaen et al (2002) by combining the 1993 household survey with the 1993 population census. The technique allows estimating consumption-based measures of poverty and inequality at very low level of geographic disaggregation. Mistiaen et al (2002) document a considerable degree of spatial heterogeneity in poverty across administrative units within provinces in particular across districts.

4.1. The Definition of Treatment

As described earlier, the program is gradually phased-in each year to the intervention districts. In August and September each year, the provincial directors (of Seecaline) operate a bidding session in which partner NGOs are selected. These partner NGOs are responsible for opening new sites for each district. There exists a grace period between the time when the contract (between the Seecaline and the NGOs) is signed and the time when a site is actually opened and operated because it usually takes two to four months to select sites, elect a community nutrition worker, and train the NGO personnel and the community nutrition worker.

[16] The Census covers 1385 (out of a total of 1394) communes in the country in 2001. The missing communes could not be reached because of the local security reasons. The questionnaire was administered to a focus group composed of residents of the commune.

[17] The Communes were introduces in 1995, replacing Firaisanas as the smallest administrative unit. In order to create communes, some of the Firaisanas were subsequently split or changed some of the boundaries. Our unit of analysis is the community: we are able to assign all communities to old Firaisanas/new Communes with the help of a geographic mapping provided to us by the Statistical Institute.

Because we focus on quantifying the impact of different program participation durations on the malnutrition rate of young children, we first define the length of exposure to the program. A site is observed as soon as it begins its operation. We take the date of the first weighing session at a site as the first period in which the site begins its operation, t_0, and we will call this the entry date. At any given point in time, *t*, after the site is open, the length of exposure to the program for each site *s* is defined as $d_{st} \equiv t - t_{s0}$. As shown in Figure 2, the date of entry is clustered around the first and forth quarters throughout the entire program period in which we have observations (i.e. between 1999 and 2002).

This gradual phase-in feature of the program at a roughly one-year interval allows us to construct two alternative definitions of treatment:

1. Being exposed to the program for *<u>one extra period (year</u>*): $d_t + 1$ versus d_t;[18]
2. Being exposed to the program for *<u>two extra years</u>:* $d_t + 2$ versus d_t.[19]

In our framework, the treatment indicator compares sites participating in Seecaline for a longer duration (the "treated" sites), $d_t + \tau, \tau = 1, 2$, with sites participating for a shorter duration (the "untreated" or "comparison" sites), d_t. It is important to note that d_t is time-varying and is measured at different points in time, as the impact is computed at five different quarters. As time progresses, all of the sites in our sample increase their program exposure over time (d_t), while the absolute *difference* in the length of exposure between the "treated" and "untreated" sites remains constant (τ, at one or two years intervals).

When measuring the impact <u>at a particular point in time</u>, we are able to quantify the *return to the length of exposure*. When keeping fixed the absolute difference in the length of exposure between the treated and the

[18]. We suppress the subscript *s* in d_{st}. It is possible that some of the "untreated or comparison" sites (i.e. sites with shorter exposure) have not been exposed to the program (i.e. $d_t=0$), depending on the entry date and the date when we measure the impact.

[19] We experimented with different thresholds in separating sites into different groups, where we compare sites with *<u>at least</u>* one (but less than two) years or at least two years difference in the exposure to the program. The results are qualitatively consistent and are available from the authors. The definition of treatment used in the paper makes it easier to provide a more precise and straightforward interpretation.

comparison sites and measuring the impact at different points in time, we are able to capture the *change in the return to the length of exposure.*

4.2. Characteristics of Sites with Different Lengths of Exposure to the Program

Since we focus on quantifying the return to duration of participation by comparing sites that enter the Seecaline program at different points in time, it is important to ask the question: “How different are sites that joined at different points in time?” Table 3 provides the group mean characteristics of the sites in each of the treatment categories. The comparison of these mean characteristics suggests that the program was progressively phased-in to the areas that are more isolated (both in terms of communes being more distant to the nearest urban center and further away from main roads, as well as in terms of more isolated communities within communes), with fewer infrastructure (connection to electricity and presence of markets), poorer (as indicated by the headcount ratio), and with a higher average initial malnutrition rate (measured at the opening date of the site). The differences reiterate the point that the phase-in aspect of the program is not made at random, and it is invalid to take a simple mean comparison approach that measure the difference in outcomes across the two treatment groups and attribute the difference to the length of exposure.

Importantly, the data also shows a substantial heterogeneity across sites that joined the program within the same period. Figures 2 and 3 show that the dispersion around the mean at different dates of joining the program is large along different dimensions, such as the poverty rate and the initial malnutrition rate. The extent of the dispersion is advantageous for our purposes because it gives us scope for finding sites that are “comparable” across different duration of exposure to the program with respect to the observed characteristics.

These data sources (the monitoring data, commune census, and poverty map) provide us with the key predetermined characteristics to us model the selection of communities/sites into different lengths of exposure within the intervention districts. We have information on the socio-economic characteristics of the communes such as remoteness, existing infrastructure and weather shocks that hit the commune in the three years prior to the census. The shock information helps us identify those communes in districts added only in a second phase in 2000 due to droughts and cyclones severely hitting the country. As the sequential selection of sites was done in partnership between the communes and the NGOs and the accessibility of the sites (measured by the distance between the site and the center of the commune) is one of the main constraints in the expansion of the program,, we include in our selection model the information on the remoteness of the sites, key socio-economic characteristics of the communes of intervention, NGO characteristics, and the interactions between remoteness and the NGO characteristics.

One important concern in modeling selection into different durations is to what extent we can capture the placement that are done sequentially in communities where most needed, i.e. poorer and more malnourished sites. In this regard, we control for both the poverty rate of the commune and the initial nutrition status of the target population. We take the malnutrition rate at the first measurement of the sites as the (pre-intervention) indicator of the initial health status of the target population. Since the initial malnutrition rate is a time-varying covariate, we de-trend it by estimating a flexible year trend.

In the next section, we will re-state the question that we address in this study and describe in details the sources of identification**.**

4.3. The Target Population

The population of interest is all children living in the intervention areas who are aged between zero and three, which is the age window that has the largest risk of growth failure (Shrimpton et al 2001)[20]. We categorize the target population into three age groups (0-6 months, 7-12 months and 13-36 months) and use site as the level of observation. The separate analysis by age groups is motivated by the nature of the intervention because the feeding and child care practices vary greatly across these groups. For instance, one of the main messages of the program is to encourage exclusive breastfeeding for children six months old or younger, which comprises 14% of the target population. On the other hand, children who are seven to twelve months old (20% of the target population) enter the weaning phase, a period during which semi-solid complementary food and liquids are introduced. In this transition, children are vulnerable to losing weight because this is the period when they start to be exposed to food-born illnesses. The program emphasizes the importance of the preparation and composition of food for this age

[20] Once a child gets older than three years of age, growth retardation becomes permanent and has important long term consequences on cognitive development, schooling and productivity later in life.

group. Older children (1-3 years of age) account for 64% of the target population in the participating communities (table 2). During this last age window, children are vulnerable to infections. During these critical period the immune system is still not fully mature (Martorell 1999) and susceptible to illnesses. The program emphasizes the importance of hygiene practices as well as promotes child care practices in the event of illnesses.

4.4. The Outcome of Interest

Our outcome of interest is the average malnutrition rate, which is measured by the nutritional status of children under three, of the communities participating to the program. We use weight-for-age[21] as an indicator for the nutritional status of children. If the weight (for a given age) of a particular child falls below two (three) standard deviations from the international standard (z-scores), then the child is considered as moderately (severely) malnourished.

Figure 5a presents the non-parametric density estimates of malnutrition rates over time, separated by the three age groups.. Three points are worth noting in this figure: (1) the proportion of underweight children increases with age, (2) There is a downward trend in the outcome of interest in the participating communities, as malnutrition rates decrease over time for all age groups, (3) There is seasonal variation in malnutrition rates, which slightly rebound around the lean season (in the first and last quarters of each year), especially for the 7-12 and 13-36 months old children.

In addition, we document in figure 5b the trends in malnutrition rates over time, separated by age groups, for participating communities that have different lengths of exposure. As we stress earlier in this paper, the vertical distance between these trends cannot be entirely attributed to the difference in the duration of program exposure. It is because the vertical distance captures the difference in the average malnutrition rates between sites that entered the program in two different periods without accounting for initial differences in the selective selection into different program durations.

[21] The weight of each child attending the weighing sessions is obtained from the registry information.

5. Empirical Methodology

We are interested in measuring the effect on malnutrition rate of different durations of exposure to the Seecaline program and the heterogeneity of the program impacts with respect to some key socio-economic characteristics of the sites. Following Behrman, Cheng and Todd (2004), we use propensity-score matching to estimate the marginal effect of being exposed to the program for different durations. An attractiveness of this methodology is that the marginal program impact can be estimated using only data on program participants. Moreover, the approach does not require a model in the selection into the program and allows for selection based on unobservables (Behrman, Cheng, and Todd (2004)). The marginal impact is estimated by comparing "comparable" participating communities that enter the Seecaline program at different points in time. Knowing that communities are selected into the program in a non-random fashion, we estimate a model of program exposure duration selection. We focus on the aggregate outcomes at the community level rather than household-level outcomes. As the socio-economic characteristics of the participants (and non-participants) in the communities are not available in the data set, we control for aggregate socio-economic variables at low level of geographical disaggregation (i.e. communities and communes). We also exploit the panel dimension of the monitoring data to compute the differential effect <u>at different points in time</u>, as the length of exposure increases.

5.1. Estimating the differential returns to the length of program exposure

To formally define the parameter of interest, we need some notations. We first define the malnutrition rate for age group a living in site *s* and participating in the Seecaline program for d_{st} periods as of time *t* as a random variable $Y_{st}(a, d_{st})$. The differential treatment effect is defined as:

$$DTE_{st}(a, d_{st}, \tau) = E[Y_{st}(a, d_{st} + \tau) - Y_{st}(a, d_{st}) \mid d_{st} + \tau]$$

where $DTE_{st}(a, d_{st}, \tau)$ denotes the difference in the average malnutrition rate measured at time period *t* for the children in age group *a* living in the sites that have participated for $d_{st} + \tau$ years relative to the counterfactual situation in which the sites only participate for $d_{s't}$ years. In what follows, we subsume d_{st} and the subscript *s* in the differential treatment effect for notational brevity. τ measures the additional years of program exposure, where

$\tau = 1, 2$ in our case. In a nutshell, what we do is to take two groups of sites with different lengths of exposures at time period *t* and compare their average malnutrition rates, controlling for all the covariates that might affect their assignments into these different program durations. Given that the outcome of interest is malnutrition rates, $DTE_t(a,\tau) < 0$ should be interpreted as a marginal *improvement* of the nutritional status of children of age group a (in sites exposed for τ years more). Rearranging the above expression as:

$$E[Y_t(a, d_t + \tau) \mid d_t + \tau] - E[Y_t(a, d_t) \mid d_t + \tau].$$

it becomes clear that this parameter can never be directly observed in the data, for the second term represents a counterfactual outcome[22] that cannot be observed[23].

We employ the method of propensity-score matching to non-parametrically estimate the $DTE_t(a,\tau)$. This estimation strategy can produce a consistent estimator for the above parameter of interest, provided that the two identification assumptions for propensity-score matching hold. First, the counterfactual average malnutrition rates measured in period *t* of the longer-exposed sites had they instead have shorter durations of exposure can be estimated using the period *t* average malnutrition rates of the shorter-exposed sites, conditional on a vector of pre-program characteristics of the sites. The identification assumption is that the selection into different lengths of exposure is conditionally exogenous, once we adjust for differences in the exogenous background characteristics that have been used in the phase-in of the program. The second condition requires that there is sufficient overlapping support in the distribution of covariates of the two treatment groups (long and short exposure).

Using the panel structure of our data, we can quantify $DTE_t(a,\tau)$ at different points in time. Note that $DTE_t(a,\tau)$ is time-varying: as time progresses, duration d_t increases for both 'treated' and 'non-treated' sites the absolute difference in τ stays the same. Keeping in mind that $DTE_t(a,\tau) < 0$ implies a differential *improvement*, we propose two hypotheses:

[22] Specifically, the counterfactual is expresses as the average malnutrition rate at time *t* for the longer exposed ($d_t + \tau$) sites had they been exposed for a shorter duration (d_t).

[23] We also estimated the parameter of interest using a regression framework in which all of the matching variables included in the propensity-score model are included. We found that the results are even stronger in the regression framework than in the non-parametric technique. Nonetheless, we chose the non-parametric method because we want to impose in our estimation strategy the overlapping condition (comparability) across sites with different durations and do not want to impose the linear functional form on the impact of interest.

(1) $DTE_t(a,\tau) = E[Y_t(a,d_t+\tau) - Y_t(a,d_t) \mid d_t+\tau] \leq 0 \quad \forall \tau > 0$ That is, the returns to exposure are *non-negative*. Our data allows us to look at the cases where $\tau = 2$ and $\tau = 1$. We cannot map out the entire function of the returns to exposure because we only have two values of τ ; nevertheless, we can draw inferences on the slope of the function.

(2) $\Delta DTE_{t-t'}(a,\tau) = DTE_t(a,\tau) - DTE_{t'}(a,\tau) \geq 0 \quad \forall t > t'$, That is, the returns are *decreasing (non-increasing)* with the length of program exposure.

There are several concerns when estimating the differential impact of the length of program exposure. One important concern is to be able to disentangle *time effect* from the *program effect*. There might be an underlying trend for the health outcome even in the absence of the intervention due to a general trend in economic growth and welfare of the population (and therefore a downward trend in the level of malnutrition rate). Nonetheless, the existence of this trend does not automatically bias our results, insofar as the trend applies equally to all communities that participate in the program for different durations. We will assume that all participating sites experience the same trend, if any, so such a trend can be eliminated by taking the difference between sites with longer and shorter durations at any given point in time.

A related concern is to be able to separate *seasonal effect* (therefore deviations from a general trend) from the differential exposure to the program. We address the issue by comparing the estimated differential returns across the same quarter in two different years. Should the results be driven only by seasonality, we would find that the observed differential effects to be the same.

Finally, an important caveat is that the observed outcome is obtained only from the children who attended the weighing sessions. Given that our outcome measures are aggregate in nature, we cannot observe who attended (the weighing sessions) and who did not. Bias may arise from the possibility that the weighed children represent systematically the more (or less) malnourished children in the community. In an appendix, we provide a sensitivity analysis and show that our results are not affected by the non-compliance of some of the eligible participants in the participating communities.

5.2. Heterogeneity in returns and its policy implications

Besides quantifying the returns to the length of program exposure, we assess the effectiveness of a large-scale intervention and draw policy implications from it. Unlike small-scale interventions, which often focus on a fairly homogeneous group of participants, large-scale, but not necessary universal, programs bring about heterogeneity in characteristics (and possibly returns) across different regions and areas of a particular country. The impact estimates of pilot program cannot be generalized to a larger population or even to the same population in different areas of the country (what is commonly referred to in the evaluation literature as a problem of *external validity*).[24]

In the context of our analysis, we estimate how the (differential) returns to the program vary according to some key aggregate economic and environmental factors that affect the child health production function: poverty levels, incidence of cyclones, the length of the lean season and access to drinkable water. Poverty levels proxy for the average level of economic conditions and resources/nutrients availability. Cyclones, the length of the lean season, and the lack of access to safe water are all factors that enhance the likelihood of morbidity and the likelihood of growth faltering. Promoting behavioral change in feeding and child care practices in environmental settings with low incomes and where the existing supplies of food and water are contaminated might offer sizable benefits to a child's health. We estimate whether the marginal gains for "participating in Seecaline earlier" vary systematically along these dimensions.

In order to measure how the DTE varies according to each of these dimensions of interest, we divide the sample in two subgroups (with/without cyclones/water, with poverty or length of the lean season below/above the median) for each of the variables ($X_{t_0}^{k}$) that capture these characteristics. We then re-estimate the *DTE* for each of these subgroups, i.e. we estimate $DTE_t(a, \tau, X^{-k}; X^k = j)$ for two different groups $j = 0,1$.

Should the returns to exposures be systematically correlated with some socio-economic characteristics of the areas of intervention, the results could provide guidance for future expansions of the program into different areas of the country.

24 Another important effect coming with scale are general equilibrium effects, generally defined as cases where the intervention itself on a large scale affect both demand, supply, and the equilibrium price (Moffitt 2004, Heckman, Lalonde and Smith 1999). In our case, the intervention promotes behavioral changes, and it does not affect directly a factor or product market (as in the case, for example of a labor market intervention).

This extrapolation to other settings and populations can be done only under clear assumptions (Heckman, Vytlacil 2006): (i) The results hold only for a given support of the distribution of environmental variables $F(X_{t_o})$ and can be extrapolated only to other areas that have the same support on these socio-economic characteristics. (ii) Under our identification assumption, selection into different lengths of exposure is exogenous, conditional on the observable characteristics X_{t_0}, 'treatment' is independent to unobservable. In order to extrapolate the results of the DTE to other areas, one would have to assume that the relationship between observables and unobservables are the same across the current intervention areas and the new potential target populations.

6. Results

The first step to estimate the return to program duration of exposure is to estimate the probability of a site being exposed for differential durations, as a function of the key observable factors that were used in the program site selection process. The estimated marginal effects of propensity score model (presented in table 4) summarize the main features of how the program was phased-in over time. Among these features, accessibility and the characteristics of the NGOs play the most important role. Sites located in more accessible communes and in more accessible villages within communes were chosen to participate in the program earlier. These 'early' sites were also relatively "better-off" as indicated by the coefficient estimate on the poverty rate of the commune. The NGO characteristics and their interaction with the remoteness of the sites are jointly significant, and the sites operated by larger NGOs and the NGOs working on environment or agriculture sectors were chosen to start earlier, and relatively more so in more accessible sites. On the other hand, the NGOs with more experience in the district before their partnership with the program began with the relatively more remote areas. The weather shock variables are jointly significant, reflecting the timing of addition of new districts for cyclones and droughts in 1999-2000.

6.1 Returns to the Length of Program Exposure

Table 5 reports the main results of the estimated treatment impact of the length of program exposure on the malnutrition status of children in three different age groups, where the impacts are measured at five different quarters. The returns to one year of additional exposure are positive and significant for the younger age groups.

Put it differently, living in sites that participate in Seecaline for one extra year, the 0-6 months old children experience an average of 8% reduction in malnutrition rate while the corresponding figure is 4% for the 7-12 months old children. When the extra year of program exposure increases to two, children in all three age groups enjoy an average of 8% reduction in malnutrition rate. With an average initial malnutrition rate of 46 percentage points, an average of eight percentage points reduction in malnutrition rates for all age groups amounts to around 20% of the initial malnutrition rate. These results are even more substantial for the younger children, whose average initial malnutrition rate is lower to begin with. In sum, this large-scale community-based nutrition program exert a positive and (statistically and economically) significant impact on all three age groups for the two lengths of program exposure.

The gains to program exposure for the 13-36 months old children materialize only after two extra years of exposure. Sites that have two additional years of exposure have impacts that are comparable across all three age groups. We interpret this pattern as a reflection of the *composition effect*: With shorter duration of program exposure, younger children, relatively to their older counterpart, live a longer fraction of their lives in sites that are covered by the program. As sites continue their participation in the program, even the older children get exposed to the program for a longer fraction of their lives. Put it differently, it takes the 0-6 months old children six months to receive the "full dose" of treatment, defined as the situation where all children live their entire lives under the program. Therefore, when living only in sites that expose for one extra year, the majority of the children in the 13-36 months old group have lived their lives without the program.

Another notable feature of the results is that the estimated program impact decreases over time. This decrease in estimated impact is confirmed by looking at the same quarter in different years, which eliminates the potential confounding effect of seasonality of malnutrition rates. A possible explanation is the "*catching-up*" argument: Recall that the monitoring data only collects information on the participating sites. Although both the treated and potential comparison sites always remain one and two years of difference in the length of exposure to the Seecaline program, all of these sites increase their program participation durations as we follow them over time. As time progresses, sites with a shorter duration "*catch up*" (at least partially) to a level of nutritional status closer to that of sites with the longer duration. As the potential comparison sites also started to receive the program services, the nutritional status of their children will improve.

In our data, we observe that this estimated impact is decreasing over time but does not converge to zero, even for those age groups who are fully exposed. This finding is consistent with the hypothesis that the return to the length of exposure decreases as the length of exposure increases; that is, the health production function is concave in the length of exposure.

6.2 Impact Heterogeneity

After investigating the overall impact of the program, we examine how the program impact varies with some of the important aggregate environmental factors underlying the health production function. These factors include poverty, incidence of cyclones, and the length of the lean season. Table 6 reports their pair-wise correlations while table 7 presents the results of the estimated DTE for the subgroups of each of the covariate of interest.

Sites with a higher poverty rate have higher returns to exposure over two years for all age groups relatively to sites with lower poverty rates (panel 7a). Panels 7b-d of table 7 present all of the estimates related to variables affecting morbidity. Sites that are attacked more by cyclones and sites that experience a longer lean season have higher DTE. The systematic higher returns are evident over durations of extra two years, especially for children aged 7-12 and 13-36, who are the age groups that are more susceptible to diseases. The results on accessibility to safe water, however, do not exhibit any significant impact, but we suspect that our measure of access to water in the data is not perfect: It applies only to the presence of public wells[25] only in the main village within the commune, so the actual access to piped water could be quite heterogeneous across villages within the same commune.

Taken together, these subgroup analyses show that the differential returns to exposure are higher in areas with scarcer resources and areas that are more vulnerable to diseases. The marginal benefits from having (earlier) the

[25] More specifically, the variable access to safe drinkable water is equal to 1 if the commune has any piped water provided by the national or by a local water utility. There is also very limited information on the relative quality of water provided by the national water utility with respect to the local providers.

Seecaline program in the economically disadvantaged environment suggests that if a large-scale nutrition intervention is designed to phase in over time, then it pays to target the needier communities first.

7. Conclusion and Discussion

This paper quantifies the impact of a large scale community-based nutritional intervention targeted at the children below the age of three, lactating women, and pregnant women implemented in Madagascar since 1999. The contribution of the study is to show how to community-level administrative data that is collected as part of the regular process of monitoring program activities can be used for the purpose of documenting the impact of the program. Such exercise can have high potential payoffs in terms of applications to other contexts, since most social programs in developing countries are carefully monitored by a monitoring system. The data produced during this process has two key attractive features that are not shared by household level data: ‘scale’ and ‘time’. Scale implies that monitoring data does represent the entire universe of participants and therefore can be potentially used to characterize the heterogeneity of impact. The time dimension enriches the analysis to be implemented at different point in times, documenting the time evolution of impact. This type of data however comes with its own limitations. First, monitoring data (and administrative data in general) do not generally have a large set of covariates. We argue that this disadvantage can be overcome by complementing these data with census-based data with socio-economic information available at low levels of geographical disaggregation. Second, by its nature monitoring data is observed only for participating communities. This does not allow the researcher to estimate the counterfactual of no participation. However, we argue that monitoring data can be used to define a new parameter of interest that measures the *differential returns to exposure to the program*.

This parameter is estimated by comparing the malnutrition rates of communities that were exposed to the program for a longer duration with (comparable) communities that were exposed to the program for a shorter period of time. With the data at hand the comparison can be done at different points in time. We exploit the phase-in features of the program to model the selection of communities into different program durations and use the method of matching to estimate the differential effects of the length of exposure on aggregate community malnutrition rates. Our results show that on average there are positive returns to exposure: communities that were exposed to the program for an extra one (or two) years displayed lower malnutrition rates than the communities that had shorter length of exposure. The reduction in malnutrition rates for different age groups takes time to materialize, and it takes up to two years for all of the three age groups (0-6 months, 7-12 months, 13-36 months) to achieve a 7-9 percentage points reduction in malnutrition rates. With an initial malnutrition rate of around 46 percentage points, this effect on malnutrition rates amounts to 15 to 20 percentage points over a two year period. When looking at the time profile of these effects, we find that the differential returns to exposure decrease over time although they do not converge to zero as time and duration increase.

Furthermore, when examining how the effects of differential length of exposure to the program vary according to some key social and environmental conditions, we find that the differential returns are higher for poorer communities, and communities more vulnerable to cyclones and to a longer lean season.

The results do not address the question of not whether the program should be pursued or not, but rather how it should expanded once it a small scale (internally valid) evaluation has shown promising and significant effects of the program and a country has decided to pursue such an approach on a larger scale. How should the program be phased-in geographically and where are the highest marginal returns from program expansion? Our study shows that the use of monitoring data creates an opportunity for applications in the context of large-scale programs, where the socio-economic characteristics and the potential returns to the program vary greatly across different intervention areas. In this respect, the current analysis can be complemented by the evaluations that are based on household survey data. Indeed, our future research will utilize household survey data to quantify other parameters of interest and highlight how the estimated return to exposure quantified in this study is related to the estimated parameters of interest in our future studies that use the household survey data.

References

Alderman H., J., Hoddinott, and B. Kinsey (2006) "Long Term Consequences Of Early Childhood Malnutrition", *Oxford Economic Papers*, 58: 450–74.

Allen L. and S.Gillespie (2001) "What Works? A Review of Efficacy and Effectiveness of Nutrition Interventions", Asian Development Bank, Nutrition and Development Series # 5. http://www.ifpri.org/pubs/books/whatworks/whatworks.pdf

Armecin G., et al (2005) "Early Childhood Development Programs and Children's Development: Evidence from the Philippines", mimeo, Department of Economics, University of Pennsylvania.

Attanasio, O., C. Meghir and M. Székely (2003) "Using randomized experiments and structural models for 'scaling up': evidence from the PROGRESA evaluation", IFS Working Papers, EWP04/03.

Behrman J. (1996) "The Impact of Health and Nutrition on Education", *World Bank Research Observer*, 11(1): 23-39.

Behrman J. R., Y. Cheng and P. Todd (2004) "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach", *Review of Economics and Statistics*, 86(1): 108-132.

Berg, A. (1987) *Malnutrition: What can be done?*, The World Bank, Washington D.C.

Cebu Study Team (1992) "Child Health Production Function", *Journal of Development Economics*, 38: 323-51.

Currie J. and D. Thomas (1995) "Does Head Start Make a Difference?" *American Economic Review*, 85(3): 341-364.

Favin, M. and M. Griffiths (1999) *Communication for Behavior Change in Nutrition Projects*, Nutrition Toolkit Module No. 9, Human Development Network, the World Bank http://www.micronutrient.org/idpas/pdf/2019CommunicationFor.pdf

Gertler, P. (2004) "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment", American Economic Review, 94(2): 336-341.

Gillespie S., J. Mason, and R. Martorell (1996) "How Nutrition Improves", ACC/SCN Nutrition Policy Discussion Paper No. 15, UN ACC/SCN, Geneva. http://www.unsystem.org/scn/archives/npp15/index.htm

Gillespie S., M. McLachlan and R. Shrimpton eds. (2003) *Combating Malnutrition. Time to Act*. World Bank, Washington DC

Grantham-McGregor S., L. Fernald and K. Sethuraman (1999) "Effects of Health and Nutrition on Cognitive and Behavioural Development in Children in the First Three Years of Life. Part 1. Low Birth weigth, Breastfeeding and Protein-energy Malnutrition", *Food and Nutrition Bullettin*, 20(1): 53-75.

Haddad L., et al (2003) "Reducing Malnutrition: How Far does Income Growth take us?", *World Bank Economic Review*, 17, pp. 107-131

Heckman, J., R. Lalonde and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", *Handbook of Labor Economics*, Volume 3, Ashenfelter, A. and D. Card, eds., Amsterdam: Elsevier Science.

Heckman, J., and E. J. Vytlacil (2006), "Econometric Evaluation of Social Programs", chapter in the *Handbook of Econometrics*, Volume 6**,** J. Heckman and E. Leamer (eds), North Holland: Amsterdam, forthcoming.

Heaver R. (2002a) "Improving Nutrition. Issues in Management and Capacity Building", HNP Discussion Paper, World Bank.

Heaver R. (2002b) "India's Tamil Nadu Nutrition Program. Lessons and Issues in Management and Capacity Development", HNP Discussions Paper, World Bank.

Heaver R. and Y, Kachondam (2002) "Thailand's National Nutrition Program. Lessons and Issues in Management and Capacity Development", HNP Discussions Paper, World Bank.

Marek, T. et al (1999), “Successful Contracting of Prevention Services: Fighting Malnutrition in Senegal and Madagascar”, *Health Policy and Planning*, 14(4): 382-89.

Moffitt, R. A. (2004) “Forecasting the Effects of Scaling Up Social Programs: An Economics Perspective”, Mimeo, Johns Hopkins University, Dept. of Economics.

OED (2005) “Maintaining Momentum towards the MDGs: An Impact Evaluation of Interventions to Improve Maternal and Child Health and Nutrition Outcomes in Bangladesh”, World Bank.

Piehl, Anne, Suzanne Cooper, Anthony Braga and David Kennedy (2003), “Testing for Structural Breaks in the Evaluation of Programs,” Review of Economics and Statistics 85(3): 550-558.

Plowman B. et al (2004) *BASICS II Evaluation of the AIN Program in Honduras, 2002*. Published by the Basic Support for Institutionalizing Child Survival Project (BASICS II) for the United States Agency for International Development, Arlington, Virginia.

Ravallion, Martin (2000) “Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved”, The World Bank Economic Review,. 14(2): 331-45.

Ross, J. (1997) “Cost-effectiveness of the Nutrition Communication Project in Mali.” SARA Project (AED): Washington, D.C.

Save the Children (2003) “Thin on Ground: Questioning the evidence behind World-Bank funded community nutrition projects in Bangladesh, Ethiopia and Uganda”, London, UK.

Shrimpton R, et al (2001) “Worldwide Timing of Growth Faltering: Implications for Nutritional Interventions”, Pediatrics, 107(5): 1-7

Strauss J. and D. Thomas (1995) “Human Resources: Empirical Modeling of Household and Family Decisions”, in J.R. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Volume IIIA. Amsterdam: North-Holland.

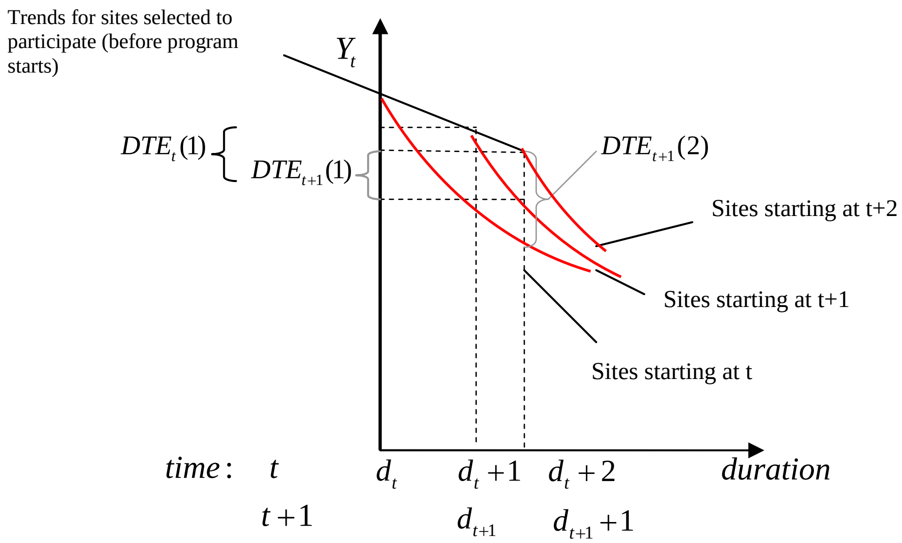
USAID (1988) “Growth Monitoring and Nutrition Education: Impact Evaluation of an Effective Applied Nutrition Program in the Dominican Republic”, CRS/CARITAS, 1983–1986. Washington, D.C.

Van Roekel K et al (2002) Mid Term Evaluation of the AIN Program in Honduras 2000, BASICS II for USAID, Arlington, Virginia.

World Health Organization (2003) *Global Strategy for Infant and Young Child Feeding*, http://www.who.int/nut/documents/ gs_infant_feeding_text_eng.pdf

Figure 1. Different Ways to Picture the Differential Effects to Program Exposure

(a) decreasing returns to exposure



(a) constant returns to exposure

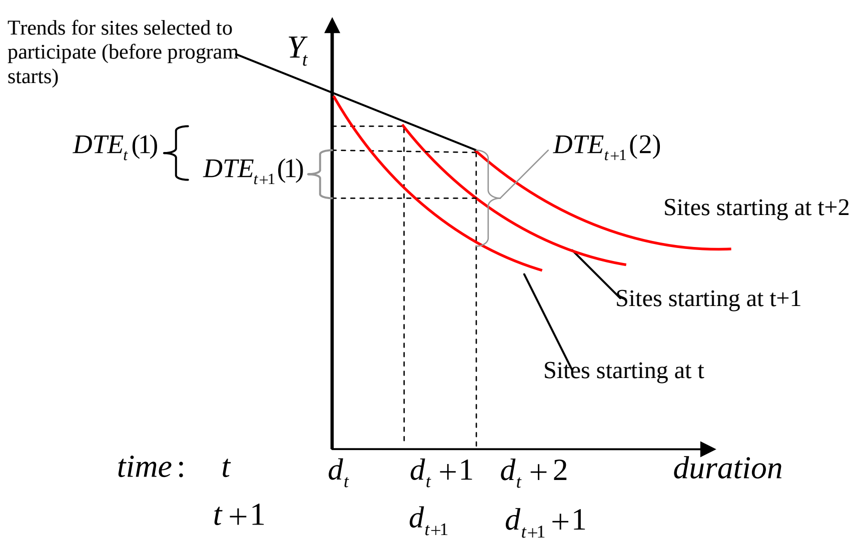


Table 1: Sample Sizes of Different Treatment Groups

	# of sites joining the program in a particular quarter	Treatment Definition: 1 more year =0	1 more year =1	2 more years =0	2 more years =1
1999q1	107				
1999q2	2				
1999q3	70				
1999q4	258				***258***
2000q1	210				***210***
2000q2	13				
2000q3	87				
2000q4	342		***342***		
2001q1	613		***613***		
2001q2	21				
2001q3	7				
2001q4	403	***403***		***403***	
2002q1	525	***525***		***525***	
2002q2	15				
2002q3	5				
2002q4	1				
Total	2697	928	955	928	468

Figure 2: Proportion of Sites that Joined the Program by date

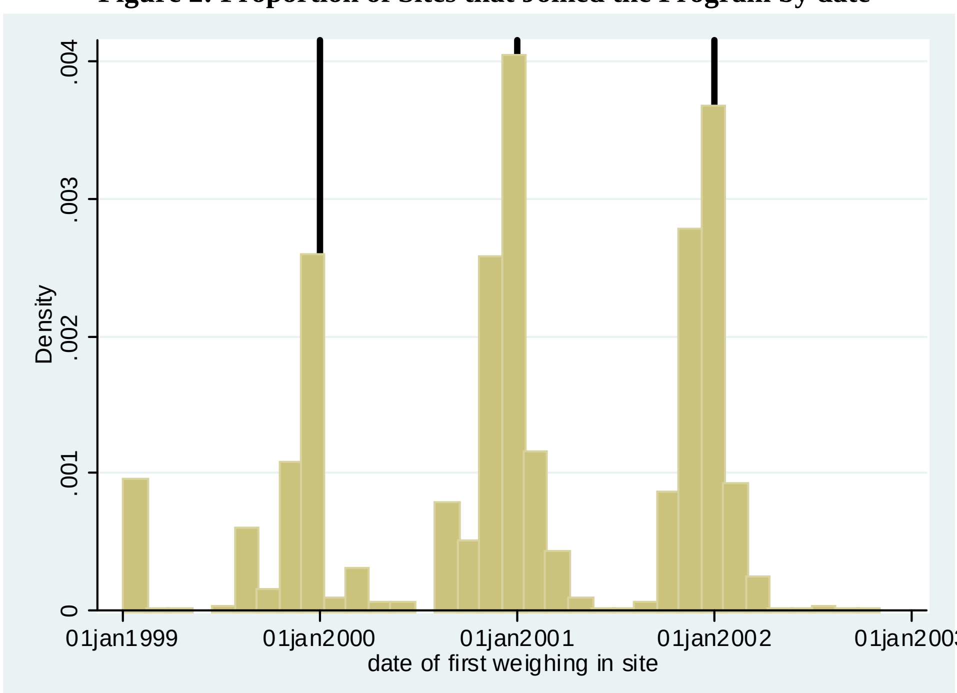


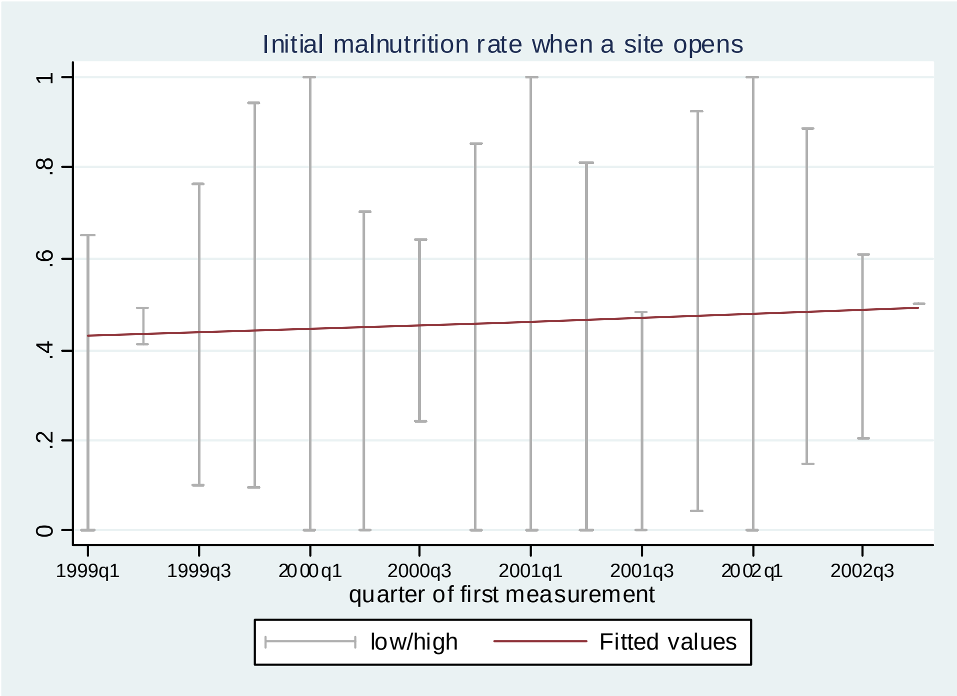
Table 2. Descriptive Statistics

Variables	Mean	S.D.	Min	Max
Community level information				
distance site from center commune (km)	9.25	10.26	0	150
initial malnutrition rate (detrended)	0.46	0.17	0.01	0.99
initial malnutrition rate (at first measurement)	0.46	0.16	0	1
initial malnutrition rate 0-6	0.31	0.23	0	1
initial malnutrition rate 7-12	0.46	0.19	0	1
initial malnutrition rate 13-36	0.51	0.18	0	1
Number of children registered	219.7	68.4	11.7	1999
Share of children registered: 0-6	0.14	0.07	0.00	0.76
Share of children registered: 7-12	0.22	0.08	0.00	0.96
Share of children registered: 13-36	0.64	0.12	0.07	1
Number of children weighed	166.9	64.5	1.5	1048
Share of children weighed: 0-6	0.14	0.08	0.00	1
Share of children weighed: 7-12	0.21	0.09	0.00	0.96
Share of children weighed: 13-36	0.66	0.13	0.01	1
Take-up rate	0.75	0.17	0.04	1
Commune level information				
distance to nearest town (km)	226.17	268.07	0	1836
presence of a national road	0.41	0.49	0	1
seasonal market	0.21	0.41	0	1
cattle market	0.31	0.46	0	1
biweekly market	0.70	0.46	0	1
daily market	0.45	0.50	0	1
log population	9.61	0.66	6.80	12.07
zone rouge	0.17	0.37	0	1
electricity	0.34	0.47	0	1
potable water	0.50	0.50		
- water provided by JIRAMA (public water utility)	0.09	0.29	0	1
- potable water provided by local providers	0.43	0.49	0	1
hospital	0.12	0.33	0	1
existence agribusiness	0.38	0.72	0	3
poverty rate (headcount ratio), 1993	0.78	0.10	0.31	0.95
cyclone in 1998-1999	0.33	0.47	0	1
flooding in 1998-1999	0.49	0.50	0	1
drought in 1998-1999	0.42	0.49	0	1
cyclone in 1999-2000	0.34	0.47	0	1
flooding in 1999-2000	0.55	0.50	0	1
drought in 1999-2000	0.38	0.48	0	1
NGO characteristics				
NGO-working in health	0.47	0.50	0	1
NGO-working in agric/environment	0.51	0.50	0	1
NGO-years of experience	6.15	6.93	0	47
NGO log personnel (net of Seecaline)	2.51	0.91	0	4.56
national NGO	0.31	0.46	0	1
Provincial NGO	0.11	0.31	0	1
local NGO	0.60	0.49	0	1

Table 3: Comparison of community characteristics according to the length of exposure

Variables	(a) Less than 1 year of exposure		(b) 1 more year		(c) 2 more years		t-test (a) vs (b)	t-test (a) vs (c)
Community level information	mean	*s.d.*	mean	*s.d.*	mean	*s.d.*		
distance site chef lieu commune (km)	10.06	*9.17*	8.44	*9.42*	8.01	*9.77*	**	**
initial malnutrition rate (detrended)	0.47	*0.17*	0.47	*0.17*	0.45	*0.17*		
Commune level information								
distance to nearest town (km)	261.85	*301.64*	261.29	*283.37*	166.83	*230.94*		**
presence of a national road	0.39	*0.49*	0.41	*0.49*	0.41	*0.49*		
seasonal market	0.19	*0.39*	0.21	*0.41*	0.25	*0.44*		**
cattle market	0.22	*0.41*	0.36	*0.48*	0.28	*0.45*	**	**
biweekly market	0.63	*0.48*	0.72	*0.45*	0.69	*0.46*	**	**
daily market	0.40	*0.49*	0.40	*0.49*	0.56	*0.50*		**
log population	9.60	*0.65*	9.59	*0.69*	9.63	*0.61*		
zone rouge	0.15	*0.36*	0.12	*0.33*	0.20	*0.40*	**	**
electricity	0.27	*0.44*	0.34	*0.47*	0.46	*0.50*	**	**
drinkable water	0.41	*0.49*	0.54	*0.50*	0.54	*0.50*	**	**
hospital	0.10	*0.30*	0.11	*0.32*	0.16	*0.37*		**
existence agribusiness	0.34	*0.71*	0.31	*0.67*	0.67	*0.86*		**
poverty rate (headcount ratio)	0.78	*0.09*	0.78	*0.11*	0.75	*0.12*		**
cyclone in 1998-1999	0.30	*0.46*	0.35	*0.48*	0.32	*0.47*	**	
inondation in 1998-1999	0.37	*0.48*	0.32	*0.47*	0.34	*0.48*	**	
drought in 1998-1999	0.07	*0.25*	0.08	*0.27*	0.09	*0.28*		
cyclone in 1999-2000	0.38	*0.49*	0.47	*0.50*	0.36	*0.48*	**	
inondation in 1999-2000	0.35	*0.48*	0.43	*0.50*	0.32	*0.47*	**	
drought in 1999-2000	0.28	*0.45*	0.33	*0.47*	0.15	*0.36*	**	**
NGO characteristics								
local NGO	0.56	*0.50*	0.67	*0.47*	0.62	*0.49*	**	**
NGO-working in health	0.49	*0.50*	0.37	*0.48*	0.59	*0.49*	**	
NGO log personnel (net of Seecaline)	2.45	*0.97*	2.59	*0.90*	2.45	*0.85*	**	

Figure 3: Phase-in and Initial Malnutrition Rate of the Sites



Each vertical represents the range of malnutrition rate of the sites that entered the program in a particular quarter.

Figure 4: Phase-in According to the Poverty Level of the Commune

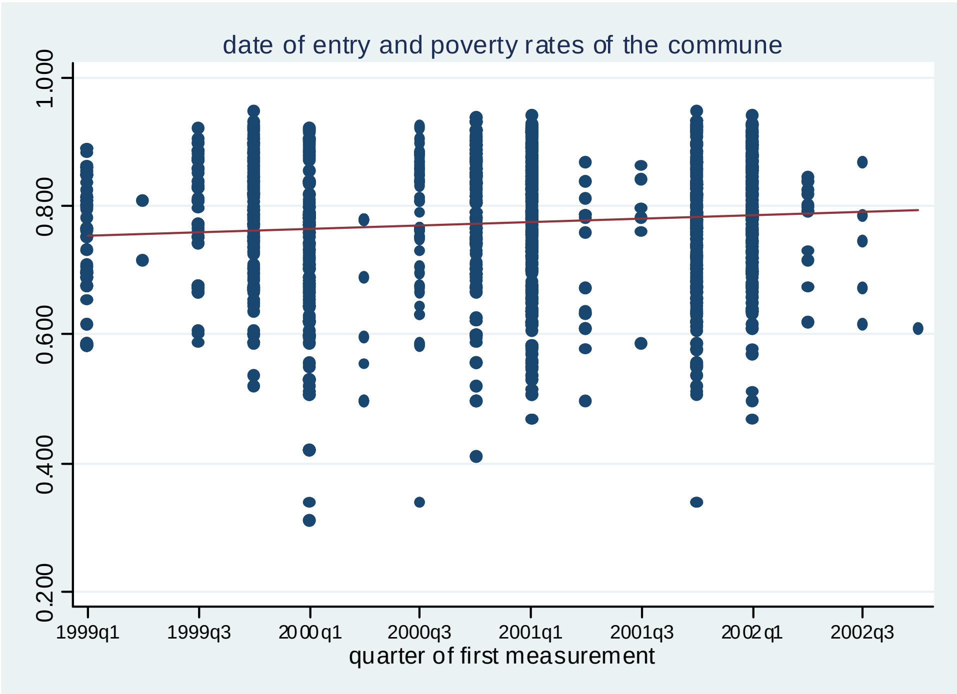


Figure 5a: Trends in the Malnutrition Rates between 2000 and 2002

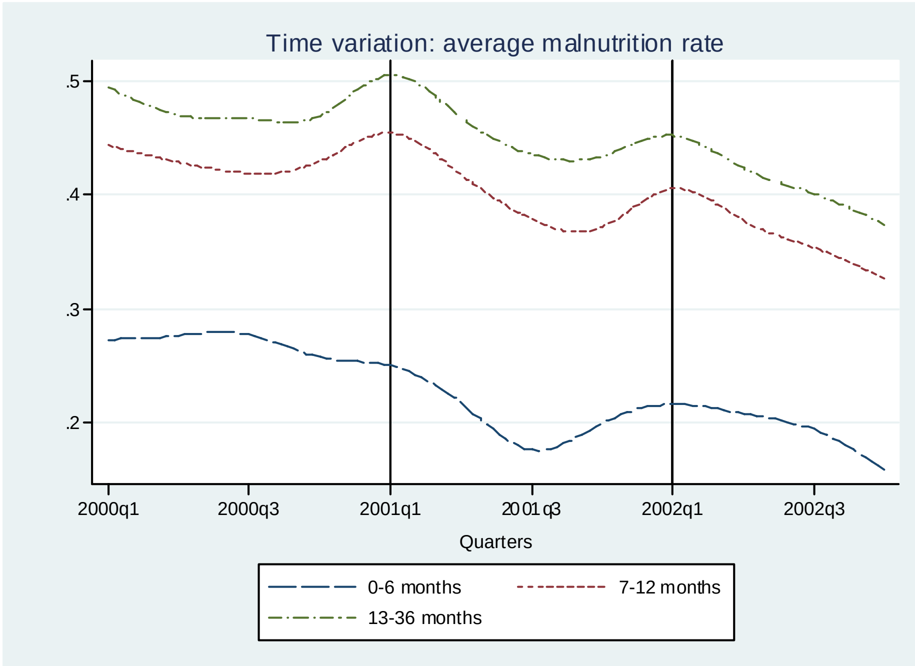


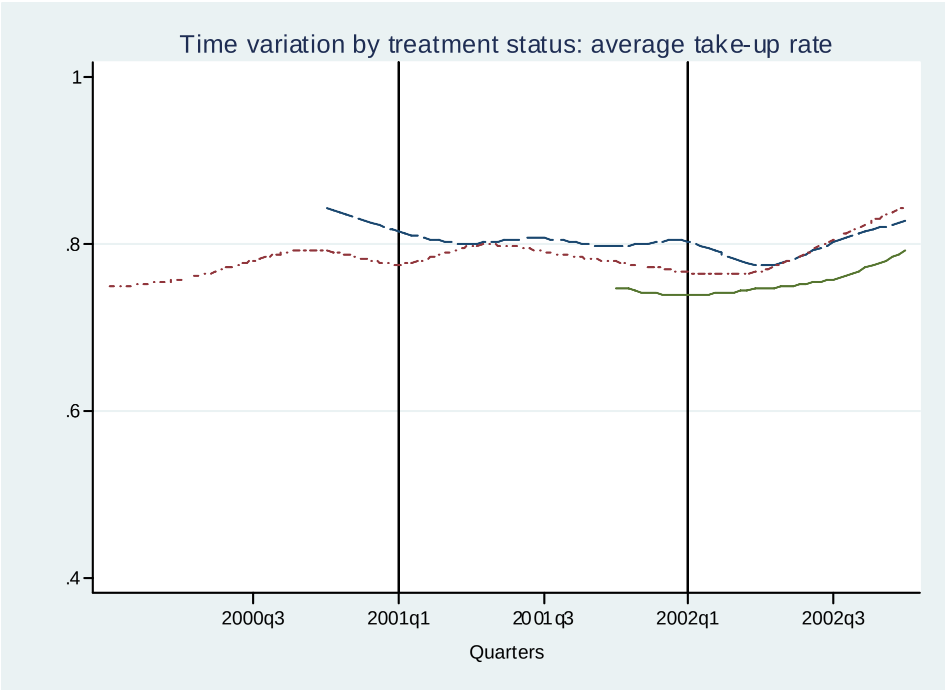
Figure 5b: Trends in the Malnutrition Rates by treatment status

Note: red line: sites exposed for 2 more years, blue line: sites exposed for 1 more year, green line: sites exposed for less than one year.

Figure 6a: Trends in Average Take-up Rates by Age Groups

Note: Take-up rate is defined as the proportion of all eligible children who participated to the weighing session at a given point in time

Fig. 6b. Trends in the take-up rates by treatment status



Note: red line: sites exposed for 2 more years, blue line: sites exposed for 1 more year, green line: sites exposed for less than one year.

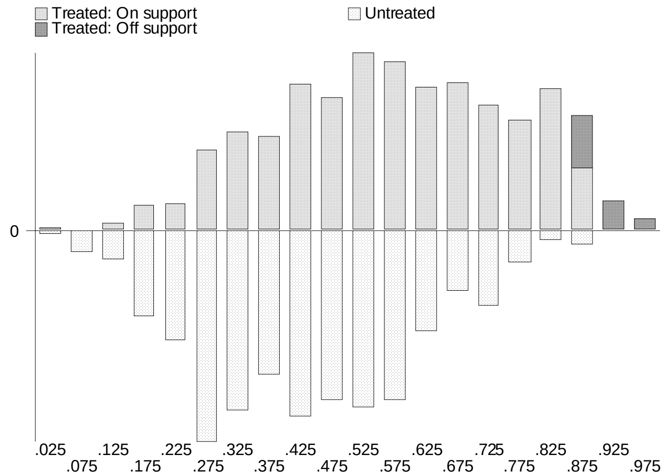
Table 4. Propensity Score estimates

Explanatory Variables	T=1 extra year exposure marginal effect	t-stat	T=2 extra years exposure marginal effect	t-stat
distance site-commune (km)	0.004	0.83	-0.016	-1.91
distance commune nearest urban center (km)	0.000	0.92	0.000	-1.76
proximity national road	-0.039	-1.29	-0.025	-0.78
seasonal market	0.060	1.76	0.087*	2.33
cattle market	0.041	1.14	-0.048	-1.31
bi-weekly market	0.046	1.56	0.044	1.37
log(population)	0.009	0.40	-0.018	-0.71
Zonerouge (insecurity zone)	-0.116*	-3.00	0.033	0.79
electricity	0.096*	2.68	0.089*	2.41
access to potable water	0.122*	4.24	-0.010	-0.31
hospital	-0.131*	-2.58	0.153*	2.94
main activity: manufacturing	-0.055*	-2.47	0.034	1.64
poverty rate (FGT0)	-0.417*	-2.93	-0.150	-1.01
local NGO	0.067	1.69	-0.097*	-2.23
NGO working in health	-0.033	-0.83	0.039	0.96
NGO working in agr./environ.	0.250*	6.33	0.129*	3.07
NGO: years of experience	-0.003	-0.84	0.000	-0.09
log(personnel NGO)	0.076*	3.57	0.048*	2.00
initial malnutrition rate site	0.035	0.42	-0.048	-0.48
Distance site *NGO local	0.001	0.40	0.016*	4.18
Distance site *NGO health	-0.007	-1.91	0.005	1.50
Distance site *NGO agriculture	-0.013*	-3.97	-0.008*	-2.24
distance site *NGO experience	0.001*	3.20	0.001*	3.24
Distance site *NGO log(pers)	-0.004*	-2.54	-0.002	-0.79
cyclone in 1999	0.010	0.31	-0.044	-1.31
flooding in 1999	0.021	0.68	0.024	0.66
drought in 1999	0.067*	2.18	0.016	0.49
cyclone in 2000	0.058	1.67	-0.052	-1.49
flooding in 2000	-0.101*	-2.87	0.016	0.43
drought in 2000	0.013	0.42	0.013	0.39
Province: Fianarantsoa	0.080	1.66	-0.184*	-3.86
Province: Tamatave	0.111*	2.56	-0.229*	-5.73
Province: Tulear	0.326*	6.24	-0.114	-1.87
Number of Observations	1802		1332	
Observed Probability	0.507		0.333	
Predicted Probability	0.511		0.305	
LR chi2	284.91		249.10	
Prob > chi2	0.000		0.000	
Pseudo R2	0.114		0.147	

Note: * indicates that the estimate is statistically significant at the 5% level. T-statistics are based on robust standard errors.

Figure 7: Distributions of the Estimated Propensity Scores

Treatment: 1 extra year



Treatment: two extra years

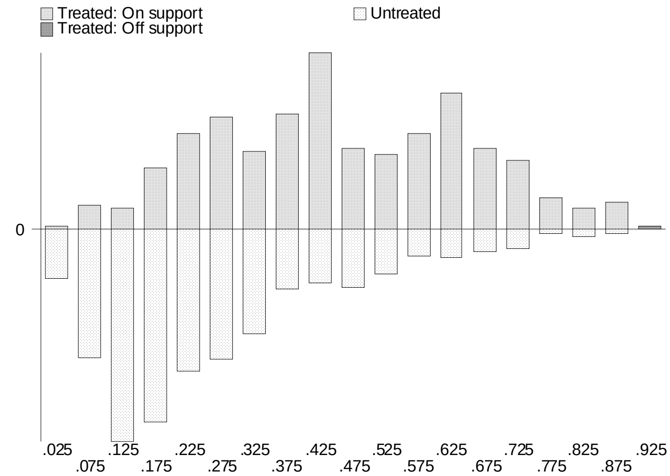


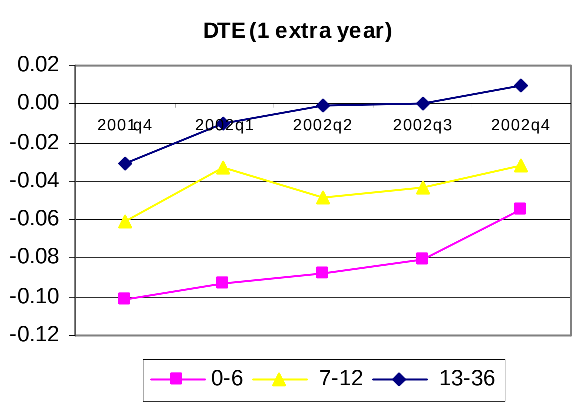
Table 5: Differential Treatment Impact, by age group

	Being exposed to SEECALINE for one more year $(\tau = 1)$			**Being exposed to SEECALINE for 2 more years** $(\tau = 2)$		
	0-6 months	7-12 months	13-36 months	0-6 months	7-12 months	13-36 months
2001Q4	-0.10*	-0.06*	-0.03*	-0.09*	-0.10*	-0.09*
	[-0.13,-0.07]	[-0.08,-0.04]	[-0.05,-0.01]	[-0.12,-0.06]	[-0.12,-0.08]	[-0.12,-0.06]
2002Q1	-0.09*	-0.03*	-0.01	-0.11*	-0.10*	-0.09*
	[-0.12,-0.06]	[-0.05,-0.01]	[0.03,0.01]	[-0.14,-0.01]	[-0.12,-0.07]	[-0.11,-0.06]
2002Q2	-0.09*	-0.05*	0.00	-0.10*	-0.11*	-0.10*
	[-0.11,-0.06]	[-0.05,-0.02]	[-0.02,0.03]	[-0.13,-0.08]	[-0.13,-0.09]	[-0.12,-0.05]
2002Q3	-0.08*	-0.04*	0.00	-0.08*	-0.07*	-0.08*
	[-0.087,-0.05]	[-0.05,-0.02]	[-0.01,0.02]	[-0.10,-0.07]	[-0.09,-0.05]	[-0.10,-0.05]
2002Q4	-0.05*	-0.03*	0.01	-0.06*	-0.04*	-0.04*
	[-0.08,-0.03]	[-0.05,-0.01]	[-.008,.039]	[-0.08,-0.04]	[-0.07,-0.02]	[-0.05,-0.02]
2002 (Avg)	-0.08	-0.04	0.00	-0.09	-0.08	-0.08

Note: 1. * represents statistically significance at the 5% level.

2. The range of numbers in the square brackets measures the bootstrapped confidence interval.

Fig. 8 Differential treatment effects



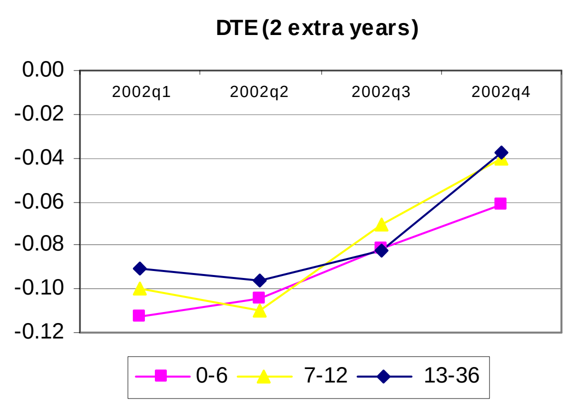


Table 6. Basic correlation coefficients among selected socio-economic characteristics

	Initial maln. Rate (community)	Poverty rate (commune)	Cyclones (commune)	Length lean season (commune)	Access to safe water (commune)
Initial malnutrition rate	1.00				
Poverty rate (headcount)	0.10	1.00			
Being exposed to cyclones	0.04	-0.01	1.00		
Length (months) lean season	0.02	0.04	0.01	1.00	
Access to safe water	0.04	-0.13	0.03	-0.09	1.00

Table 7: Differential Treatment Effects, by sub-groups

Panel 7a: Subgroups according to the poverty rate of the commune

Treatment: Being exposed to SEECALINE for *one more year* $(\tau = 1)$

	Low Poverty Sites			**High Poverty Sites**			**National Sample**		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.09*	-0.07*	-0.03	-0.07*	-0.07*	-0.06*	-0.10	-0.06	-0.03
2002 Q1	-0.09*	-0.04*	0.01	-0.10*	-0.02	-0.03	-0.09	-0.03	-0.01
2002 Q2	-0.09*	-0.05*	0.01	-0.08*	-0.04*	-0.02	-0.09	-0.05	0.00
2002 Q3	-0.06*	-0.04*	0.00	-0.09*	-0.04	0.00	-0.08	-0.04	0.00
2002 Q4	-0.04*	-0.02	0.02	-0.06*	-0.03	0.01	-0.05	-0.03	0.01
2002 (Avg)	*-0.07*	*-0.05*	*0.00*	*-0.08*	*-0.04*	*-0.02*	*-0.08*	*-0.04*	*0.00*

Treatment: Being exposed to SEECALINE for *two more years* $(\tau = 2)$

	Low Poverty Sites			**High Poverty Sites**			**National Sample**		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.07*	-0.09*	-0.08*	-0.12*	-0.12*	-0.12*	-0.09	-0.10	-0.09
2002 Q1	-0.10*	-0.09*	-0.07*	-0.13*	-0.11*	-0.14*	-0.11	-0.10	-0.09
2002 Q2	-0.09*	-0.10*	-0.09*	-0.13*	-0.12*	-0.11*	-0.10	-0.11	-0.10
2002 Q3	-0.06*	-0.07*	-0.08*	-0.11*	-0.06*	-0.08*	-0.08	-0.07	-0.08
2002 Q4	-0.05*	-0.04*	-0.02*	-0.10*	-0.06*	-0.07*	-0.06	-0.04	-0.04
2002 (Avg)	*-0.07*	*-0.08*	*-0.07*	*-0.12*	*-0.09*	*-0.10*	*-0.09*	*-0.08*	*-0.08*

Note: 1. * represents statistically significance at the 5% level.
2. The range of numbers in the square brackets measures the bootstrapped confidence interval.

Low (high) poverty sites are defined as those communities residing in communes with a poverty rate below (above) the median poverty rate in the intervention areas (0.79). Note the average poverty rate for all rural communes is around 0.75.

Panel 7b: Subgroups according to whether the commune had ever been exposed to cyclones (1999-2001)

Treatment: Being exposed to SEECALINE for *exactly one more year* $(\tau = 1)$

	No-cyclone Areas			**Cyclone Areas**			**National Sample**		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.09*	-0.04*	-0.07*	-0.10*	-0.07*	-0.01	-0.10	-0.06	-0.03
2002 Q1	-0.08*	-0.05*	-0.03*	-0.11*	-0.02*	0.01	-0.09	-0.03	-0.01
2002 Q2	-0.06*	-0.04*	-0.02	-0.10*	-0.04*	0.02	-0.09	-0.05	0.00
2002 Q3	-0.06*	-0.04*	0.00	-0.09*	-0.04*	0.02	-0.08	-0.04	0.00
2002 Q4	-0.03	-0.04*	0.02	-0.06*	-0.02	0.03	-0.05	-0.03	0.01
2002 (Avg)	*-0.06*	*-0.04*	*-0.02*	*-0.10*	*-0.04*	*0.02*	*-0.08*	*-0.04*	*0.00*

Treatment: Being exposed to SEECALINE for *exactly two more years* $(\tau = 2)$

	No-cyclone Areas			**Cyclone Areas**			**National Sample**		
Age group	0-6	7-12	13-36	0-6	7-12	13-36	0-6	7-12	13-36
2001 Q4	-0.13*	-0.09*	-0.08*	-0.09*	-0.10*	-0.11*	-0.09	-0.10	-0.09
2002 Q1	-0.12*	-0.08*	-0.05*	-0.13*	-0.12*	-0.12*	-0.11	-0.10	-0.09
2002 Q2	-0.11*	-0.08*	-0.05	-0.12*	-0.13*	-0.12*	-0.10	-0.11	-0.10
2002 Q3	-0.07*	-0.04	-0.03	-0.09*	-0.07*	-0.09*	-0.08	-0.07	-0.08
2002 Q4	-0.05*	-0.02	-0.01	-0.06*	-0.04*	-0.04*	-0.06	-0.04	-0.04
2002 (Avg)	*-0.10*	*-0.06*	*-0.04*	*-0.10*	*-0.09*	*-0.10*	*-0.09*	*-0.08*	*-0.08*

Note: 1. * represents statistically significance at the 5% level.

2. The range of numbers in the square brackets measures the bootstrapped confidence interval.

Communes are defined to be in a cyclone area if it has been exposed to any cyclone in the period between 1999 and 2001. About 53% of the sample has been exposed to at least one cyclone during this period (relative to a national incidence of 46%).

Panel 7c: Subgroups according to the length (in months) of the lean season

Treatment: Being exposed to SEECALINE for exactly one more year $(\tau = 1)$

	Lean Season <=4months			**Lean Season >4months**			**National Sample**		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.10*	-0.07*	-0.02	-0.10*	-0.06*	-0.03*	-0.10	-0.06	-0.03
2002 Q1	-0.09*	-0.04	0.01	-0.10*	-0.03*	-0.01	-0.09	-0.03	-0.01
2002 Q2	-0.09*	-0.05*	0.01	-0.08*	-0.04*	0.00	-0.09	-0.05	0.00
2002 Q3	-0.06*	-0.04	0.00	-0.08*	-0.04*	0.00	-0.08	-0.04	0.00
2002 Q4	-0.04*	-0.02	0.02	-0.05*	-0.03	0.01	-0.05	-0.03	0.01
Average	*-0.07*	*-0.04*	*0.01*	*-0.08*	*-0.03*	*0.00*	*-0.08*	*-0.04*	*0.00*

Treatment: Being exposed to SEECALINE for exactly two more years $(\tau = 2)$

	Lean Season <=4months			**Lean Season >4months**			**National Sample**		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.05*	-0.07*	-0.09*	-0.18*	-0.14*	-0.10*	-0.09	-0.10	-0.09
2002 Q1	-0.09*	-0.09*	-0.09*	-0.16*	-0.11*	-0.09*	-0.11	-0.10	-0.09
2002 Q2	-0.08*	-0.10*	-0.11*	-0.16*	-0.12*	-0.08*	-0.10	-0.11	-0.10
2002 Q3	-0.06*	-0.06*	-0.08*	-0.11*	-0.09*	-0.07*	-0.08	-0.07	-0.08
2002 Q4	-0.06*	-0.03	-0.03	-0.07*	-0.06*	-0.04	-0.06	-0.04	-0.04
2002 (Avg)	*-0.10*	*-0.06*	*-0.04*	*-0.10*	*-0.09*	*-0.10*	*-0.09*	*-0.08*	*-0.08*

Note: 1. * represents statistically significance at the 5% level.
2. The range of numbers in the square brackets measures the bootstrapped confidence interval.
About 44% of the sample has a lean season above the median length of 4 months per year.

Panel 7d: Subgroups according to whether the commune had access to safe water

Treatment: Being exposed to SEECALINE for *exactly one more year* $(\tau = 1)$

	with access			without access			National Sample		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2001 Q4	-0.11*	-0.08*	-0.05*	-0.10*	-0.08*	-0.03*	-0.10	-0.06	-0.03
2002 Q1	-0.11*	-0.05*	-0.01	-0.09*	-0.04*	-0.02*	-0.09	-0.03	-0.01
2002 Q2	-0.08*	-0.06*	-0.01	-0.10*	-0.04*	-0.01	-0.09	-0.05	0.00
2002 Q3	-0.08*	-0.06*	0.01	-0.10*	-0.05*	-0.03*	-0.08	-0.04	0.00
2002 Q4	-0.04*	-0.03*	0.02	-0.08*	-0.05*	-0.02	-0.05	-0.03	0.01
2002 (Avg)	*-0.08*	*-0.05*	*0.00*	*-0.09*	*-0.05*	*-0.02*	*-0.08*	*-0.04*	*0.00*

Treatment: Being exposed to SEECALINE for *exactly two more years* $(\tau = 2)$

	with access			without access			National Sample		
Age group	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*	*0-6*	*7-12*	*13-36*
2002 Q1	-0.13*	-0.12*	-0.13*	-0.06	-0.09*	-0.06*	-0.09	-0.10	-0.09
2002 Q2	-0.10*	-0.09*	-0.09*	-0.14*	-0.10*	-0.06*	-0.11	-0.10	-0.09
2002 Q3	-0.10*	-0.10*	-0.08*	-0.14*	-0.12*	-0.07*	-0.10	-0.11	-0.10
2002 Q4	-0.07*	-0.08*	-0.06	-0.13*	-0.09*	-0.08*	-0.08	-0.07	-0.08
2002 (Avg)	*-0.08*	*-0.08*	*-0.07*	*-0.13*	*-0.09*	*-0.06*	*-0.09*	*-0.08*	*-0.08*

Note: 1. * represents statistically significance at the 5% level.
2. The range of numbers in the square brackets measures the bootstrapped confidence interval.
About half of the sample of sites resides in a commune with access to drinkable water (provided by the national water utility or a local provider (see table 2).

Appendix

Concerns for selective take-up: Sensitivity analysis

In this section, we perform sensitivity analysis to examine the robustness of our results with respect to take-up rates. The nature of our data is such that we observe the outcome of interest for each site only when children show up at the weighing sessions. Figures 6a shows that the average takeup rates over time for all age groups remain quite flat at around 80%, with a slight upward trend picking up in 2002. The same pattern is observed in Fig. 6b, when the trend of takeup rates over time is plotted by treatment status. The figure suggests that the proportion of children weighted does not monotonically vary by treatment status.

The concern for a bias arises from the fact that the weighed children might be the more (or less) malnourished children in the community. Given that our data is aggregate in nature, we cannot observe the socio-economic characteristics of those who show up at the weighing sessions (or for that matter observe those who do not participate), so that we cannot model the self-selection into the program. This selective bias could also be exacerbated by the fact that take-up rates be systematically related to length of exposure and duration. The descriptive evidence does not seem to be supportive of this link. But in principle, the correlation would arise if it takes time to mobilize communities around the program (takeup positively related to exposure) or if there is fatigue of the communities after some time (takeup negatively related to exposure).

More formally, let that $Y_{st}(a,d_{st})$ is the observed and $Y^*_{st}(a,d_{st})$ is the true malnutrition rate of the population of all mothers/children in community *s*. Let $T_{st} \in (0,1]$ be the take-up rate of age group *a* of site *s* at time t. Then, the true malnutrition rate can be expressed as:

$Y^*_{st}(a,d_{st}) - Y_{st}(a,d_{st}) = (1-k(T_{st}))$ where $k(T_{st}) > 0$ measures the extent to which the observed malnutrition rate deviates from the true one. Ideally, we would like to have the bias to be zero, which would imply $k(T_{st}) = 1$. We allow this deviation to be a function of the take-up rate. The bias in the estimated impact arising from selective take-up can therefore be expressed as:

$$\begin{aligned} Bias &= [Y_{st}(a,d_{st}+\tau) - Y^*_{st}(a,d_{st}+\tau)] - [Y_{st}(a,d_{st}) - Y^*_{st}(a,d_{st})] \\ &= \{[1-k(T_{st}(d_{st}+\tau))]Y_{st}(a,d_{st}+\tau)\} - \{[1-k(T_{st}(d_{st}))]Y_{st}(a,d_{st})\} \end{aligned}$$

If $k(T_{st})$ differs from one, then we have to consider the following two cases in order to access the sign of a systematic bias in the estimated impact

(i) If the participants are the relatively healthier children, then $Y_{st}(d_{st}+\tau) < Y^*_{st}(d_{st}+\tau)$ and $Y_{st}(d_{st}) < Y^*_{st}(d_{st})$, so that $k(T_{st}) < 1$.

(ii) If the participants are the relatively less healthy children, then $Y_{st}(d_t+\tau) > Y^*_{st}(d_t+\tau)$ and $Y_{st}(d_t) > Y^*_{st}(d_t)$ and $k(T_{st}) > 1$.

In addition, if take-up increases with duration, then $k(T_{st}(d_t+\tau))$ would be closer to one than $k(T_{st}(d_t))$ would.

Given that our descriptive statistics suggest that take-up is not related to length of exposure, (k is constant), we will work on the maintained assumption that the bias is constant.

Our inference on the presence of such bias relies on the comparison of these results for the same age group and duration by re-doing the analysis on subsamples of the data with progressively higher thresholds for the take-up rates. We first exclude sites in the bottom 25% of the (age-specific) take-up rates, and subsequently exclude the bottom 50% of such distribution. Comparing these results to the results obtained using the whole sample provides an indirect evidence of the existence of bias. In particular, to the extent that there is a systematic and monotonic bias due to the selective take-up of the program, increasing the threshold monotonically would provide an indirect inference on the direction of the bias ($k < / > 1$): if the effects under these alternative samples are systematically lower (or higher) than the one obtained in the entire sample, we would infer that we are over- (or under)estimating the differential treatment effects.

The results from this sensitivity analysis are summarized in table 8 and represented graphically in graph 7. The evidence does not suggest the presence of any systematic bias for the age groups 7-12 and 13-36. However, it seems to suggest a slight over-estimation of the estimated effect for 0-6 months' old children. Nevertheless, the extent of the bias does not seem to be quantitatively large: the point estimates of the DTE in the restricted samples falls always within the confidence interval obtained in the whole sample.

Table A1: Sensitivity analysis: increasing thresholds for age-specific take-up rates

Panel A: entire sample (replicated from table 3)

	Being exposed to SEECALINE for one more year $(\tau = 1)$			Being exposed to SEECALINE for 2 more years $(\tau = 2)$		
	0-6 months	7-12 months	13-36 months	0-6 months	7-12 months	13-36 months
2001Q4	-0.10*	-0.06*	-0.03*	-0.09*	-0.10*	-0.09*
	[-0.13,-0.07]	[-0.08,-0.04]	[-0.05,-0.01]	[-0.12,-0.06]	[-0.12,-0.08]	[-0.12,-0.06]
2002Q1	-0.09*	-0.03*	-0.01	-0.11*	-0.10*	-0.09*
	[-0.12,-0.06]	[-0.05,-0.01]	[0.03,0.01]	[-0.14,-0.01]	[-0.12,-0.07]	[-0.11,-0.06]
2002Q2	-0.09*	-0.05*	0.00	-0.10*	-0.11*	-0.10*
	[-0.11,-0.06]	[-0.05,-0.02]	[-0.02,0.03]	[-0.13,-0.08]	[-0.13,-0.09]	[-0.12,-0.05]
2002Q3	-0.08*	-0.04*	0.00	-0.08*	-0.07*	-0.08*
	[-0.08,-0.05]	[-0.05,-0.02]	[-0.01,0.02]	[-0.10,-0.07]	[-0.09,-0.05]	[-0.10,-0.05]
2002Q4	-0.05*	-0.03*	0.01	-0.06*	-0.04*	-0.04*
	[-0.08,-0.03]	[-0.05,-0.01]	[-.008,.039]	[-0.08,-0.04]	[-0.07,-0.02]	[-0.05,-0.02]
2002 (Avg)	-0.08	-0.04	0.00	-0.09	-0.08	-0.08

Panel B: Sample excluding bottom 25% take-up distribution

	Being exposed to SEECALINE for one more year $(\tau = 1)$			Being exposed to SEECALINE for 2 more years $(\tau = 2)$		
	0-6 months (take-up>0.58)	7-12 months (take-up>0.62)	13-36 months (take-up>0.66)	0-6 months (take-up>0.58)	7-12 months (take-up>0.62)	13-36 months (take-up>0.66)
2001Q4	-0.10	-0.07	-0.04			
2002Q1	-0.10	-0.04	-0.01	-0.12	-0.11	-0.10
2002Q2	-0.08	-0.05	0.00	-0.11	-0.12	-0.10
2002Q3	-0.07	-0.04	0.00	-0.07	-0.07	-0.09
2002Q4	-0.04	-0.02	0.01	-0.06	-0.04	-0.04
2002 (Avg)	-0.07	-0.04	0.00	-0.09	-0.08	-0.08

Panel C: Excluding bottom 50% take-up distribution

	Being exposed to SEECALINE for one more year $(\tau = 1)$			Being exposed to SEECALINE for 2 more years $(\tau = 2)$		
	0-6 months (take-up>0.76)	7-12 months (take-up>0.78)	13-36 months (take-up>0.81)	0-6 months (take-up>0.76)	7-12 months (take-up>0.78)	13-36 months (take-up>0.81)
2001Q4	-0.07	-0.07	-0.06			
2002Q1	-0.07	-0.04	-0.02	-0.09	-0.11	-0.13
2002Q2	-0.08	-0.04	-0.01	-0.07	-0.11	-0.13
2002Q3	-0.06	-0.04	0.01	-0.06	-0.06	-0.09
2002Q4	0.01	-0.02	0.01	-0.05	-0.05	-0.03
2002 (Avg)	-0.05	-0.04	0.00	-0.07	-0.08	-0.09

Fig. A1 Differential treatment effects: samples with increasing thresholds of take-up